

Device Health Estimation by Combining Contextual Control Information with Sensor Data

Tomonori Honda and **Linxia Liao** and **Hoda Eldardiry** and **Bhaskar Saha** and **Rui Abreu**

Palo Alto Research Center, Palo Alto, California, USA

e-mail: {tomo.honda, linxia.liao, hoda.eldardiry, bhaskar.saha, rui.maranhao}@parc.com

Radu Pavel and **Jonathan C. Iverson**

TechSolve, Inc., Cincinnati, Ohio, USA

e-mail: {pavel, iverson}@TechSolve.org

Abstract

The goal of this work is to bridge the gap between business decision making and real-time factory data. Beyond real-time data collection, we aim to provide analysis capability to obtain insights from the data and converting the learnings into actionable recommendations. We focus on analyzing device health conditions and propose a data fusion method that combines sensor data with limited diagnostic signals with the device's operating context. We propose a segmentation algorithm that provides a temporal representation of the device's operation context, which is combined with sensor data to facilitate device health estimation. Sensor data is decomposed into features by time-domain and frequency-domain analysis. Principal component analysis (PCA) is used to project the high-dimensional feature space into a low-dimensional space followed by a linear discriminant analysis (LDA) to search the optimal separation among different device health conditions. Our industrial experimental results show that by combining device operating context with sensor data, our proposed segmentation and PCA-LDA approach can accurately identify various device imbalance conditions even for limited sensor data which could not be used to diagnose imbalance on its own.

1 Introduction

The growing Internet of Things is predicted to connect 30 billion devices by 2020 [1]. This will bring in tremendous amounts of data and drive the innovations needed to realize the vision of Industry 4.0—cyber-physical systems monitoring physical processes, and communicating and cooperating with each other and with humans in real time. One of the key challenges to be addressed is how to analyze large amounts of data to provide useful and actionable information for businesses intelligence and decision making. In particular, to prevent unexpected downtime and its significant impact on overall equipment effectiveness (OEE) and total cost of ownership (TCO) in many industries. Continuous monitoring of equipment and early detection of incipient faults can support optimal maintenance strategies, prevent downtime, increase productivity, and reduce costs.

A significant number of anomaly detection and diagnosis methods have been proposed for machine fault detection and machine health condition estimation. Chandola et al. [2]

discusses various categories of anomaly detection technologies and their assumptions as well as their computational complexity. Several approaches such as statistical methods [3], neural network methods [4] and reliability methods [5], have been applied to detect anomalies for various types of equipment. The philosophies and techniques of monitoring and predicting machine health with the goal of improving reliability and reducing unscheduled downtime of rotary machines are presented by Lee et al. [6].

Many of these methods focus on analyzing, combining, and modeling sensor data (e.g. vibration, current, acoustics signal) to detect machine faults. One issue that remains mostly unaddressed in these methods is that they rarely consider the varying operating context of the machine. In many cases, false alarms are generated due to a change in machine operation (e.g. rotational speed) rather than a change in machine condition. A major challenge in addressing this issue is that most machine controllers are built with proprietary communication protocols, which leads to a barrier in obtaining control parameters to understand the context under which the machine is operating. Recently, the MTConnect open protocol [7] was developed to connect various legacy machines independent of the controller providers. MTConnect provides an unprecedented opportunity to monitor machine operating context in real-time. In this paper, we leverage MTConnect to diagnose machine health condition by combining sensor data with operating context information. Additionally, we investigate whether it is possible to diagnose machine health condition using less sensor data when it is combined with context information.

Prior work [8] has demonstrated that vibration data could be used for diagnosing machine imbalance fault conditions. Our study focuses on extending prior work by exploring various types of sensor and control data for diagnosing the imbalance of the machine tools.

Our contribution includes the following extensions:

- Combining control and sensor signals to improve accuracy.
- Utilizing a different set of sensor data such as temperature, power, flow, and lubricant/coolant pH.

Our hypothesis is that these advancements to prior work will aid in improving the diagnosis capability as well as reducing the cost of machine diagnostics by utilizing cheaper sensors.

2 Experimental Data

The data under study has been collected from experiments utilizing a machine tool monitoring system implemented on a horizontal machining center manufactured by Milltronic with *Fanuc Oi-MC* control. We have two main sources of data: (i) data from additional sensors installed on the machine, and (ii) data from the machine tool controller. This data has been collected using National Instrument equipment and software (LabVIEW).

The external sensors used for data collection include:

- power sensor that measures power using Hall effect,
- accelerometers that capture machine tool motion in 6 degrees of freedom,
- thermocouples that measure temperatures at 10 locations on the machine tool,
- pH sensor for detecting the pH level of the metalworking fluid, and
- flow rate sensor to measure metalworking fluid pump flow.

The second category consists of data collected from the controller. This data includes drive loads, absolute and relative positions, servo delays, and feed rate. The complete list of the components of the control data is listed in Pavel et al.[8].

Data has been collected in two sessions, one in 2009 and the other in 2010. Although the basic control signals are similar, they are offset by constant values (see Figure 1). Since the positional offset could cause a difference in the motion dynamics, we have treated them as separate data sets for this study.

3 Technical Approach

For each extension to prior work listed in Section 1, we have performed two main steps for creating appropriate diagnostics:

- Feature Extraction & Synthesis
- Model Selection

3.1 Feature Extraction & Synthesis

There are various approaches for condensing time series information into data mining features. Prior work has utilized transfer functions to map control signals to vibrational sensor data [8]. The diagnosis step is then reduced to comparing the features of transfer function-predicted vibration data and the sensor-derived vibration data. This approach makes sense when the control signal directly impacts the output variables of the machine. For motion control of machine tools, the estimated transfer function should be similar to the transfer function of the implemented control (like PI or PID). Typical vibration data features would include average, standard deviation, and maximum FFT values [9].

However, we would like to diagnose the state of machine using not only accelerometers, but also other sensors, such as temperature sensors. Since temperatures at various locations are not part of active control loops, there may not exist well defined transfer functions that can map control signals to temperature sensor data very accurately. In such cases where conventional features extracted from temperature signals are not correlated with the fault (imbalance) to a

sufficient degree. Additionally, if the associated sensors are too expensive to install, then data fusion may be applied.

There are three data fusion approaches typically used in machinery diagnostics [10; 11]—data-level fusion, feature-level fusion, and decision-level fusion. Data-level fusion involves combining sensor data before feature extraction, such that features contain information gathered from multiple sensors. Feature-level fusion involves generating features from each sensor separately, then fusing this set of features generated from all of the sensors coherently for diagnostics. Finally, decision-level fusion creates diagnostics from each sensor separately, then aggregates these diagnostics into a single diagnostic output.

The choice of the three types of data fusion methods is often application specific. In our application, we found that temperature sensor data cannot resolve imbalance conditions by itself and control signal data is too coarse-grained to aid in classifying imbalance conditions using the standard data-fusion techniques. Note that we did not focus on spindle acceleration data, which could diagnose imbalance on its own (see Subsection 4.1) since that would require retrofitting existing machine tools with new expensive sensors and data acquisition hardware. Ideally we would like to use the readily accessible control signals and data from inexpensive temperature sensors to diagnose imbalance. To achieve this goal, we proposed a different type of data fusion approach. We used the control signal to provide the contextual information for temperature sensor data. The control signal is used for the segmentation of sensor data, but does not directly map into feature vectors (see Subsection 4.2).

3.2 Model Selection

Since the data sets are statistically small and dimensionality of the data is increased by feature synthesis, the models to be used for imbalance classification need to be carefully chosen to avoid over-fitting. The high-dimensional data needs to be projected to a much smaller sub-space to prevent over-fitting¹. To accomplish this, the main techniques used in this study are Principal Component Analysis (PCA) [12] and Linear Discriminant Analysis (LDA) [13]. These techniques are based on linear coordinate transformation, which makes them more likely to under-fit and less likely to over-fit [14].

4 Results

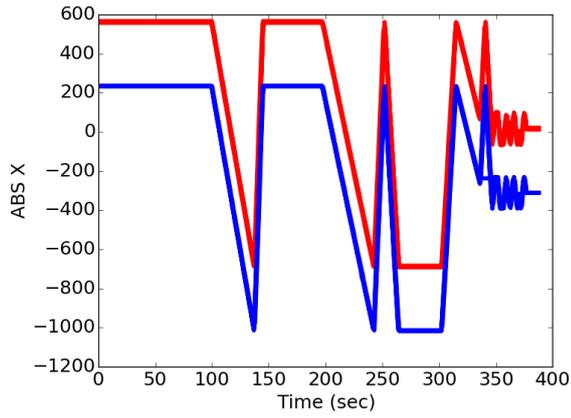
We have explored three types of imbalance diagnostics to investigate the hypothesis posed in Section 1:

- Sensor based Diagnostics
- Control based Temporal Segmentation followed by Sensor based Diagnostics

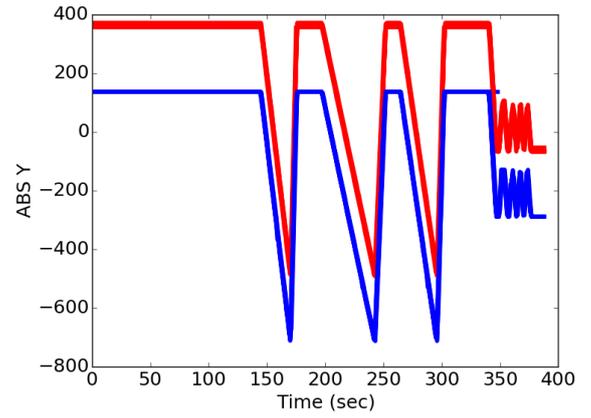
4.1 Sensor based Diagnostics

In this case, each sensor signal was analyzed separately to determine if any of the sensor signals contains enough diagnostic information to detect imbalance on its own. By plotting the time series data we find that spindle acceleration sensors (which captures vibration) show higher oscillation

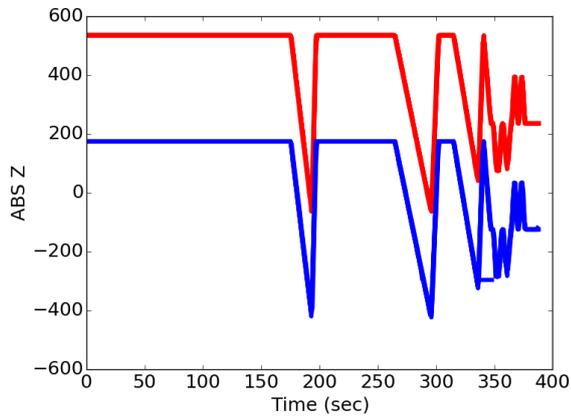
¹Note that complexity of model is positively correlated with likelihood of over-fitting. Thus, creating a classifier that takes high-dimensional input will have higher degree of freedom (i.e. higher complexity) compare to low-dimensional inputs, which results in higher likelihood of over-fitting.



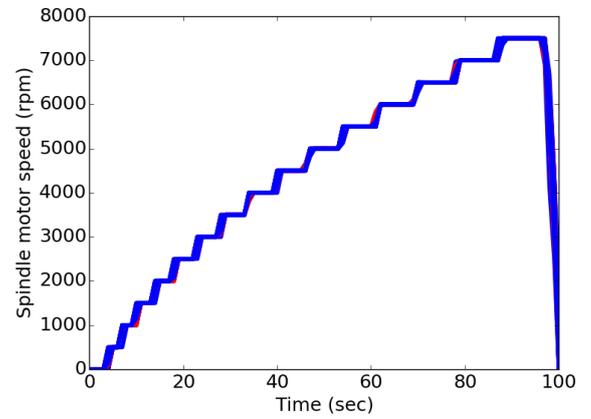
(a) Absolute X position



(b) Absolute Y position



(c) Absolute Z position



(d) Spindle Motor Speed

Figure 1: Primary Control Signals

amplitudes (see Figure 2) with increasing imbalance. Since imbalance actually impacts moment of inertia of the spindle, this change in acceleration is expected.

We also considered measuring imbalance through temperature. From the energy flow perspective, additional acceleration caused by imbalance should result in higher energy consumption from the power source and higher energy dissipation to thermal inertias due to friction, which should result in temperature increase in parts of the machine tool. However, the time series data, from each of the temperature sensors, did not show distinguishing features similar to the acceleration sensors. An example of temperature sensor time series data is shown in Figure 3.

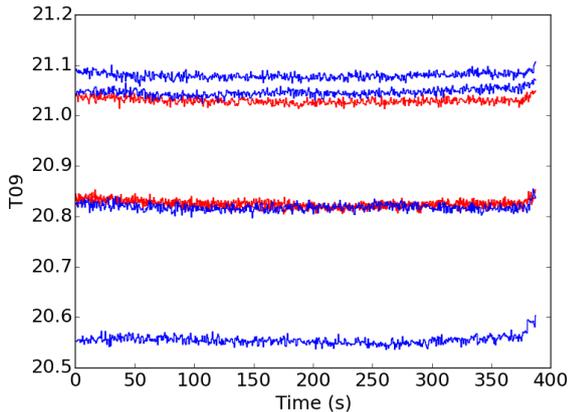


Figure 3: Sample Temperature Sensor Data (Fluid Temperature): blue and red traces indicate nominal and faulty conditions respectively

For this sensor data analysis, the features extracted are (i) average, (ii) standard deviation, (iii) maximum amplitude of FFT, and (iv) frequency for maximum amplitude of FFT. These four features are inspected visually to determine if imbalance could be classified by a simple linear classifier. The spindle acceleration (X, Y, and Z) feature (maximum amplitude of FFT) showed easily visible characteristics that can distinguish between degrees of imbalance. See Figure 4 for an example of visual classification based on X-axis acceleration data. Other sensor signals like power, pH, flow, and temperature did not exhibit such classification capability.

4.2 Control-based Segmentation followed by Sensor-based Diagnostics

The second diagnostic approach that we explored combines both sensor and control data in a coherent manner. The first step in this approach is to utilize the control signal to provide temporal segmentation, i.e., assuming quasi-steady state, the goal is to find the time intervals in which the following conditions are satisfied: (i) all experiments display same values for the primary control signal (actual spindle speed), and (ii) all the control signals are constant over the same period. Note that, to investigate the dynamic response, rather than quasi steady state response, the control signals should be consistent across the experiments so that responses are compared under the same set of control inputs. Figure 5 (a) shows the result of this temporal segmentation scheme. For each of the control signals, we have

computed the standard deviation at the each time step and identified the periods with standard deviation below a set threshold to find the consistent time intervals (shown as colored segments along the time axis in Figure 5 (b)). Then we find the intersection of the sets of consistent time intervals over all the control signals to determine the aggregate time intervals over which the control signals are statistically consistent (shown as black segments along the time axis in Figure 5 (c)).

These temporal segments are then mapped to sensor data to facilitate diagnostics. For each of 16 temporal segments, we computed features including (i) average, (ii) standard deviation, (iii) maximum FFT value, and (iv) FFT frequency at maximum amplitude. This step produces a 64 dimensional feature space to diagnose machine imbalance. As mentioned before, to avoid the overfitting we focus on linear transformation based approaches. We implemented Principal Component Analysis (PCA) to reduce the dimensionality from 64 to 4 (postulating that there should be 4 unique dimensions given the 4 uncorrelated features that we have selected). The PCA step is followed by Linear Discriminant Analysis to find the optimal coordinate transformation that provides maximum separation between classes. Result of this PCA-LDA analysis is shown in Figure 6 for Fluid Temperature sensor data. Another temperature sensor located at Spindle Motor also exhibits similar diagnostic capability after application of control based temporal segmentation. This demonstrates that control data can be used to provide context to sensor data in a way that helps diagnose machine imbalance. Thus, temperature sensor which had inferior diagnostic performance without context data, could classify imbalance perfectly when it is combined with additional context from control signal.

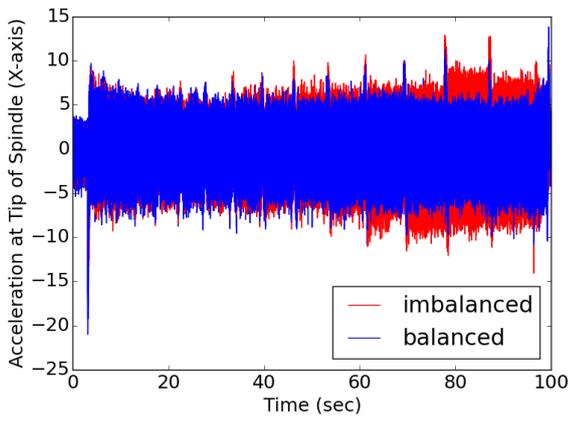
5 Conclusion and Discussion

This work explores various types of sensor and control data for diagnosing the imbalance of the machine tools. Our proposed approaches utilize sensor data that has not been used before for this purpose. This includes temperature, power, flow, and lubricant/coolant pH. In addition, our proposed techniques combine control and sensor signals to improve accuracy. Namely, by combining context information gained from the control signal, temperature sensor was able to classify machine imbalance conditions with much higher accuracy than using itself alone.

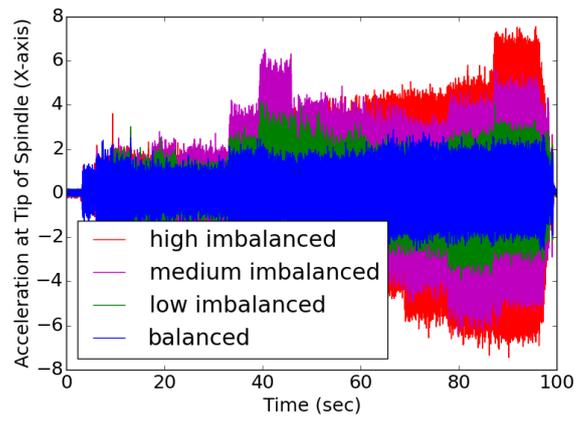
For future work, we will explore diagnostics based on control signal alone. Given that relying on sensor data typically requires adding sensors to existing machine tools, it would be ideal if we could diagnose imbalance of the machine from control signals that are usually recorded (i.e. no additional hardware required). The expectation is that if a machine tool uses feedback controls, then the control signal should be impacted by any change in the operational characteristics (in this case the imbalance of the machine tools).

References

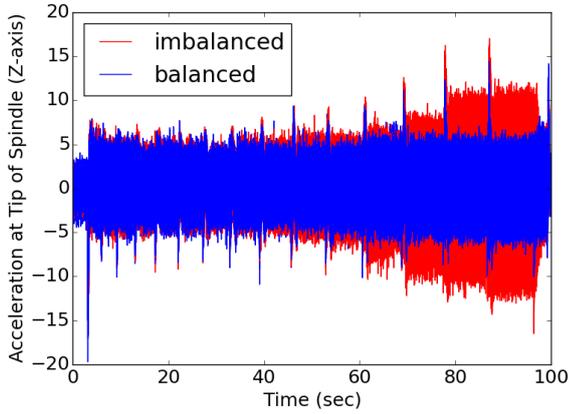
- [1] Carrie MacGillivray, Vernon Turner, and Denise Lund. Worldwide internet of things (iot) 2013–2020 forecast: Billions of things, trillions of dollars. *IDC. Doc.*, 243661(3), 2013.
- [2] Varun Chandola, Arindam Banerjee, and Vipin Kumar. Anomaly detection: A survey. *ACM Computing Surveys (CSUR)*, 41(3):15, 2009.



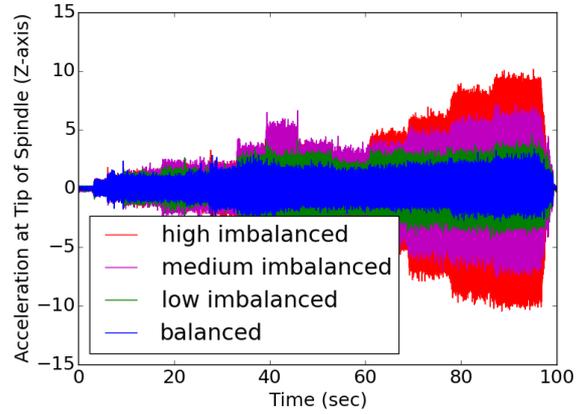
(a) Spindle X Acceleration: 2009 Data



(b) Spindle X Acceleration: 2010 Data

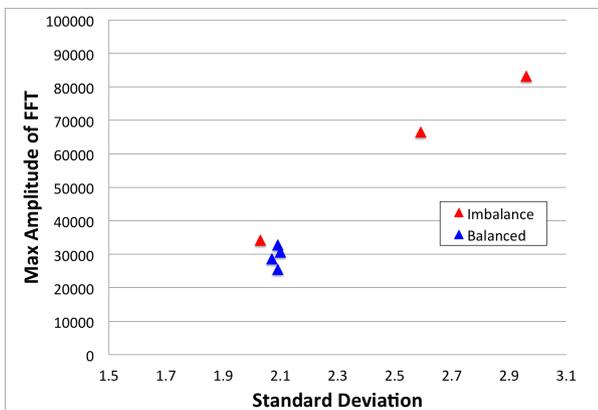


(c) Spindle Z Acceleration: 2009 Data

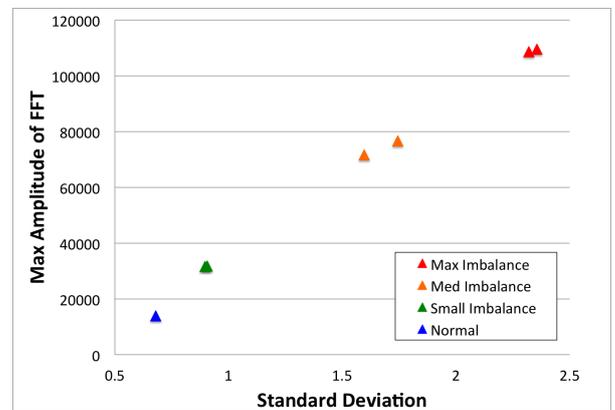


(d) Spindle Z Acceleration: 2010 Data

Figure 2: Spindle Acceleration Data for different imbalance level



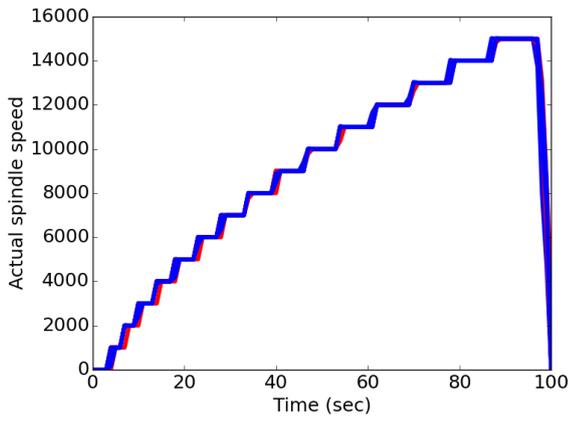
(a) Spindle X Acceleration for 2009 Data



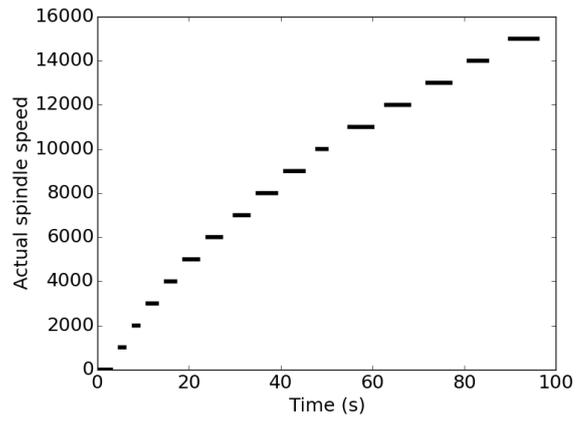
(b) Spindle X Acceleration for 2010 Data

Figure 4: Visual Classification using Spindle X Acceleration Sensor

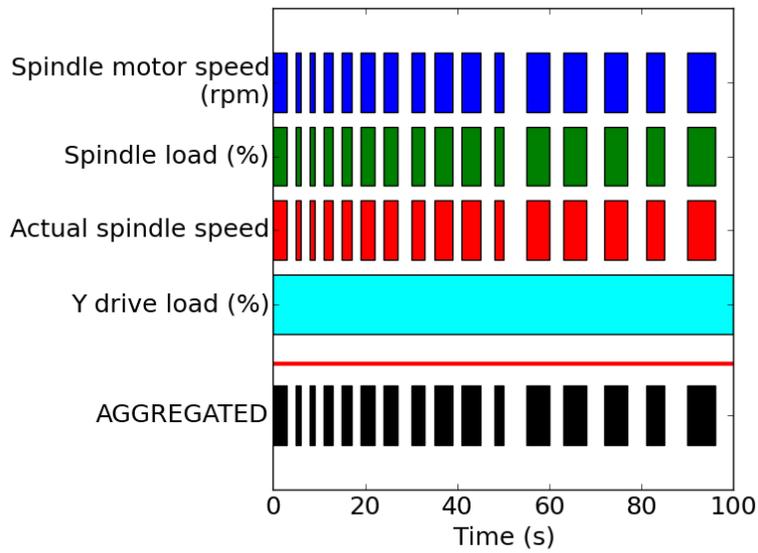
- [3] Markos Markou and Sameer Singh. Novelty detection: a review-part 1: statistical approaches. *Signal processing*, 83(12):2481–2497, 2003.
- [4] Markou Markos and Sameer Singh. Novelty detection: a review-part 2: neural network based approaches. *Signal Processing*, 83(12):2499–2521, 2003.
- [5] Haitao Guo, Simon Watson, Peter Tavner, and Jiangping Xiang. Reliability analysis for wind turbines with incomplete failure data collected from after the date of initial installation. *Reliability Engineering & System Safety*, 94(6):1057–1063, 2009.
- [6] Jay Lee, Fangji Wu, Wenyu Zhao, Masoud Ghaffari, Linxia Liao, and David Siegel. Prognostics and health management design for rotary machinery systems-reviews, methodology and applications. *Mechanical Systems and Signal Processing*, 42(1):314–334, 2014.



(a) Raw Spindle Speed Control

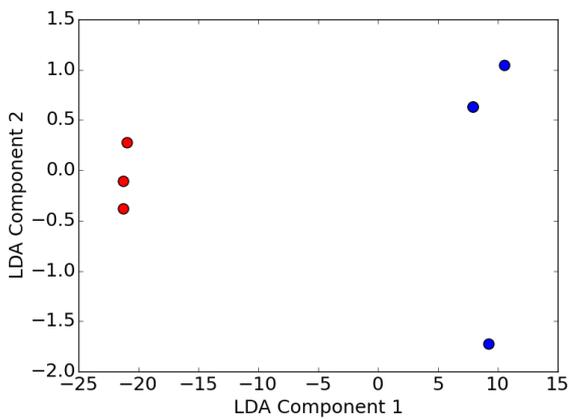


(b) Spindle Speed Control with Consistent Time Segment

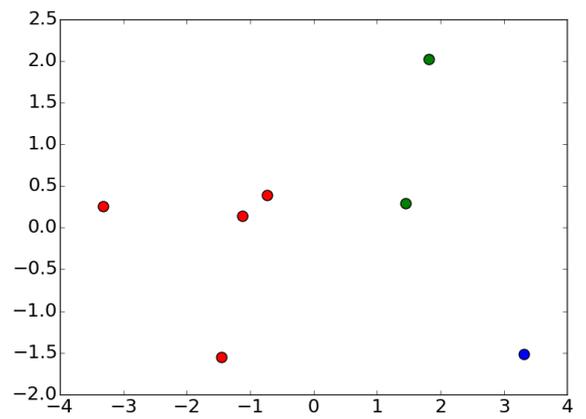


(c) Aggregating Control Signals

Figure 5: Time Series Segmentation



(a) Group 1



(b) Group 2

Figure 6: PCA-LDA Result using Fluid Temperature

[7] MTConnect Standard. Part 1-overview and protocol, version 1.01. *MTConnect Institute*, 2009.

[8] Radu Pavel, John Snyder, Nick Frankle, Gary Key, and Loran Miller. Machine tool health monitoring using

prognostic health monitoring software. In *MFPT 2010 Conference*, Huntsville, AL, April 2010.

- [9] Houtao Deng, George Runger, Eugene Tuv, and Martyanov Vladimir. A time series forest for classification and feature extraction. *Information Sciences*, 239:142–153, 2013.
- [10] Qing Charlie Liu and Hsu-Pin Ben Wang. A case study on multisensor data fusion for imbalance diagnosis of rotating machinery. *AI EDAM*, 15(03):203–210, 2001.
- [11] Andrew KS Jardine, Daming Lin, and Dragan Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing*, 20(7):1483–1510, 2006.
- [12] Svante Wold, Kim Esbensen, and Paul Geladi. Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1):37–52, 1987.
- [13] Gary J Koehler and S Selcuk Erenguc. Minimizing misclassifications in linear discriminant analysis*. *Decision sciences*, 21(1):63–85, 1990.
- [14] Bo Yang, Songcan Chen, and Xindong Wu. A structurally motivated framework for discriminant analysis. *Pattern Analysis and Applications*, 14(4):349–367, 2011.

